

Presenting the VENEC corpus: Development of a cross-cultural corpus of vocal emotion expressions and a novel method of annotating emotion appraisals

Petri Laukka¹, Hillary Anger Elfenbein², Wanda Chui³, Nutankumar S. Thingujam⁴, Frederick K. Iraki⁵, Thomas Rockstuhl⁶, & Jean Althoff⁷

¹Stockholm University, Sweden

²Washington University in St. Louis, USA

³University of California, Berkeley, USA

⁴University of Pune, India

⁵United States International University, Kenya

⁶Nanyang Technological University, Singapore

⁷University of Queensland, Australia

Recognition of affect in speech...

- ❖ ...is fast becoming one of the key disciplines in the field of affective computing
- ❖ ...is also of great relevance to both basic and applied psychological research
- ❖ *What affects/emotions can reliably be communicated through the voice?*
- ❖ *What affects/emotions are universally recognized from the voice?*

How should we annotate emotional speech?

- Most previous studies have used:
 - categorical descriptions (*e.g.*, anger, fear, joy, sadness) combined with forced choice methodology
 - dimensional ratings (*e.g.*, activation and valence)
- These methods may not capture the full spectrum of nuances that may:
 - (*a*) be present in natural conversations
 - (*b*) be perceived by listeners

Perceiving emotions = perceiving appraisals?

- Vocal expressions convey information about the speakers' emotional state, but may also reflect the antecedent appraisal processes that produced the emotion
- If vocal expressions do contain information about the emotion-antecedent evaluation processes, this "should allow the listener to reconstruct the major feature of the emotion-producing event in its effect on the speaker" (Scherer, 1988, *J Lang Soc Psychol*, p.94)

VENEC: Vocal Expressions of Nineteen Emotions across Cultures

- 100 professional actors from 5 English speaking cultures (USA, India, Kenya, Singapore, & Australia)
- 20 speakers/culture
- Ages = 18-30 years
- 50% women
- Affection, amusement, anger, contempt, disgust, distress, fear, guilt, happiness, interest, sexual lust, peacefulness, pride, relief, sadness, shame, negative surprise, positive surprise, and neutral
- 3 levels of emotion intensity

VENEC: Vocal Expressions of Nineteen Emotions across Cultures

- The actors were provided with short scenarios describing typical situations in which each emotion may be elicited, and were instructed to enact finding themselves in these situations.
- Short phrases (*e.g.*, "Let me tell you something")
- Nonlinguistic vocalizations
- Approximately 6500 stimuli with a great variety of expressive styles due to speaker, culture, and emotion intensity effects!
- Acoustically analyzed with *Praat* (\approx 70 vocal cues)

Appraisal rating study

(Laukka, Efenbein, & Chui, submitted)

- A selection from the VENEC corpus (300 stimuli)
- 20 US actors expressed 15 emotions (amusement, anger, contempt, disgust, distress, fear, guilt, happiness, negative surprise, pride, positive surprise, relief, sadness, serenity, & shame)
- 12 US listeners rated each stimulus with regard to the characteristics of the emotion eliciting situation, described in terms of appraisal dimension

Appraisal rating study

- NOVELTY - How suddenly and abruptly did the event occur? (not at all sudden/abrupt = 1; very sudden/abrupt = 5)
- INTRINSIC PLEASANTNESS - How pleasant is the event? (very unpleasant = 1; very pleasant = 5)
- GOAL CONDUCTIVENESS - Did the event help the person to reach a goal or satisfy a need? (prevented a goal/need = 1; helped a goal/need = 5)
- URGENCY - Did the person need to respond to the event urgently (not at all urgent = 1; very urgent = 5)

Appraisal rating study

- POWER - Could the outcome of the situation be modified by appropriate human action? (expressor powerless = 1; expressor powerful = 5)
- SELF-RESPONSIBILITY - Was the speaker responsible for the event? (no personal responsibility = 1; great deal of personal responsibility = 5)
- OTHER-RESPONSIBILITY - Was another person responsible for the event? (no responsibility by others = 1; great deal of responsibility by others = 5)
- NORM COMPATIBILITY - Do you think that the event was compatible with the speakers norms? (violated own norms = 1; very consistent with own norms = 5)“.

Appraisal predictions

Novelty: anger = high, fear = high, happiness = high, negative surprise = high, positive surprise = high, sadness = low, serenity = low

Pleasantness: amusement = high, anger = low, contempt = low, disgust = low, distress = low, fear = low, guilt = low, happiness = high, negative surprise = low, pride = high, positive surprise = high, relief = high, sadness = low, serenity = high, shame = low

Goal conduciveness: anger = low, distress = low, fear = low, guilt = low, happiness = high, negative surprise = low, pride = high, relief = high, sadness = low, shame = low

Urgency: amusement = low, anger = high, distress = high, fear = high, happiness = low, relief = low, sadness = low, serenity = low

Power: anger = high, contempt = high, distress = low, fear = low, guilt = low, happiness = high, pride = high, sadness = low, shame = low

Norm compatibility: amusement = high, anger = low, contempt = low, guilt = low, happiness = high, pride = high, shame = low

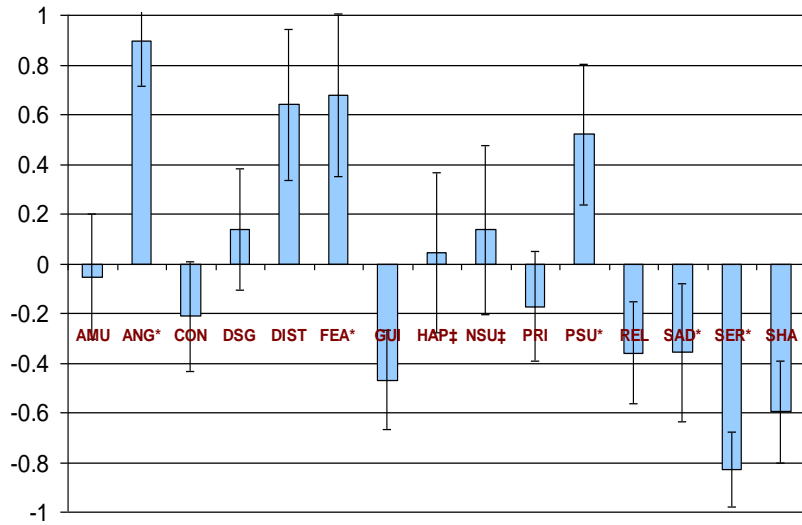
Results

- Listeners were able to rate the stimuli on appraisal scales in a consistent fashion (except for self- and other-responsibility)
- The *average measure intraclass correlation* was calculated for each scale using the Spearman-Brown formula
- Novelty ($R = 0.87$)
- Pleasantness (0.87)
- Goal conduciveness (0.85)
- Urgency (0.83)
- Power (0.81)
- Self-responsibility (0.62)
- Other-responsibility (0.52)
- Norm compatibility (0.83)

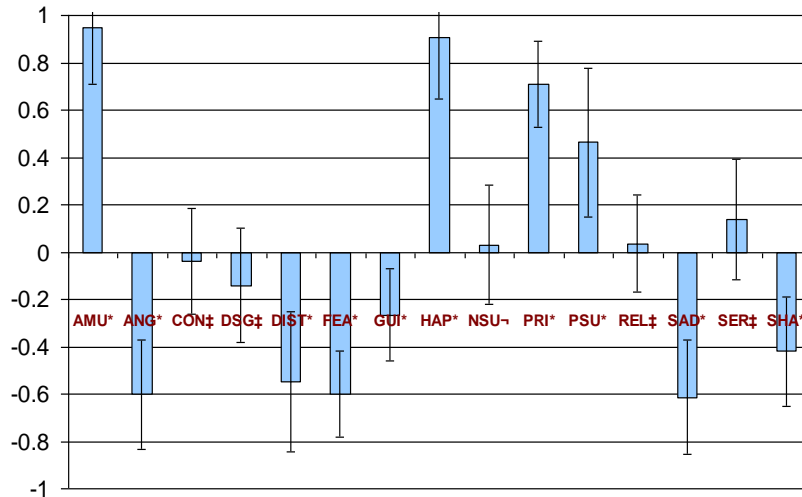
Results

- The listeners' appraisal ratings varied systematically as a function of intended emotion, as evidenced by significant effect of intended emotion in repeated measures ANOVAs conducted separately for the listeners' mean ratings on each appraisal dimension

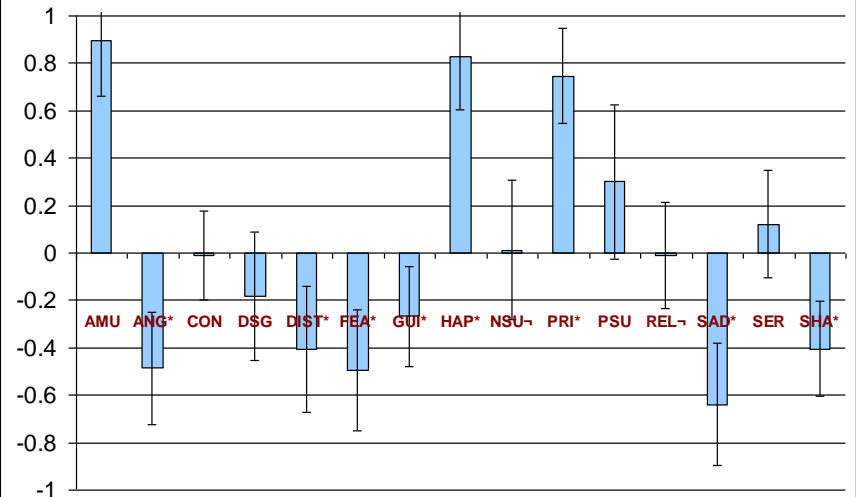
Novelty



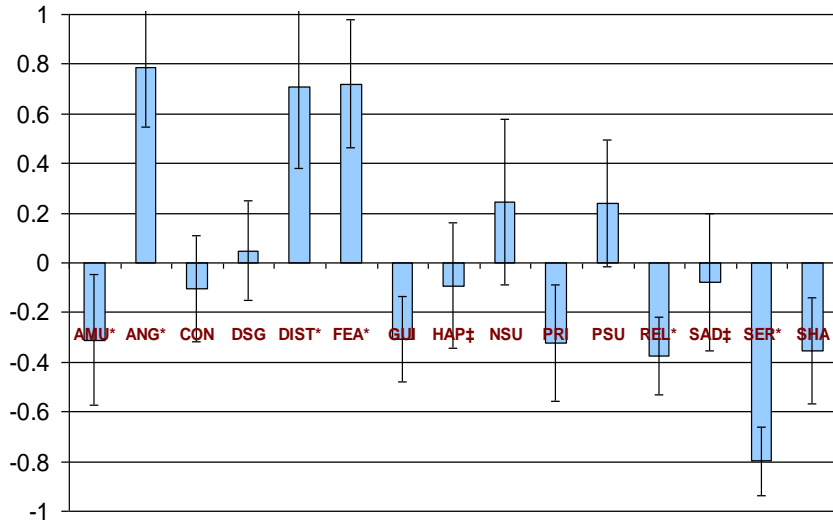
Pleasantness



Goal conduciveness

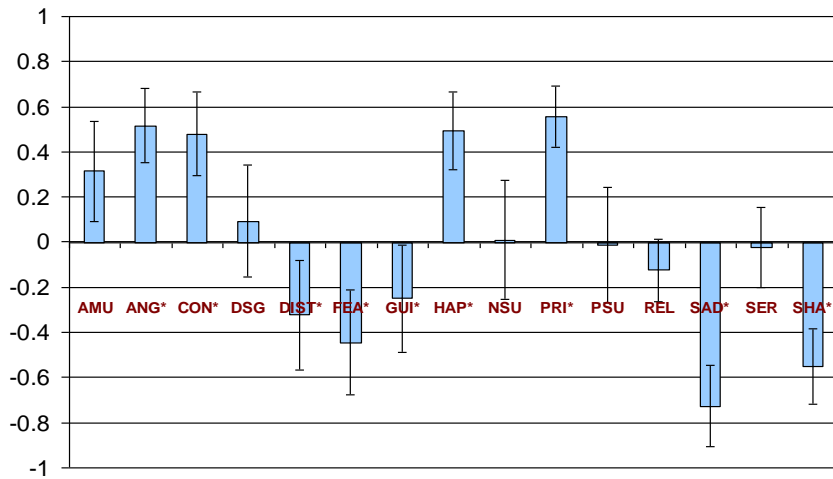


Urgency

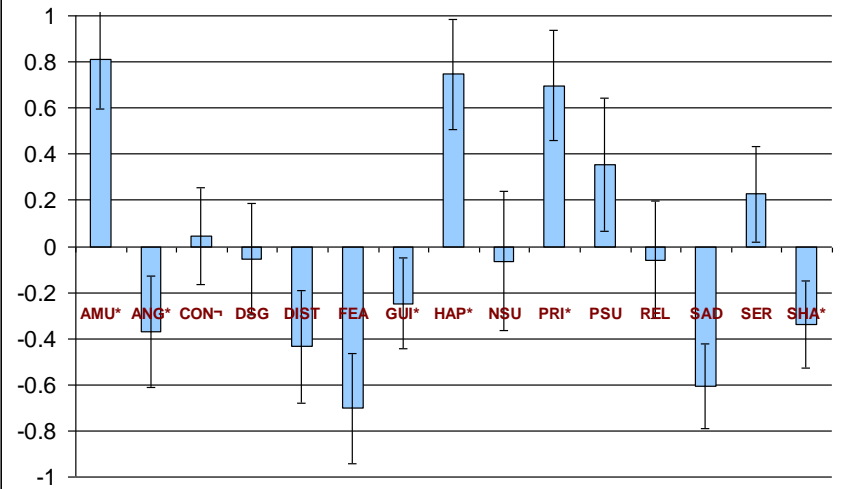


> 90% of the effects of intended emotion were in the predicted direction

Power



Norm compatibility



Acoustical correlates of appraisal dimensions

- The listeners' appraisal ratings were further significantly correlated with several acoustic variables
- Most correlations were small to medium, but give indications of which acoustic cues the listeners utilized in order to make their inferences about the emotion-eliciting situations
- We also investigated if the listeners' appraisal ratings could be predicted from the acoustic cues using multiple regression analyses. R was relatively high for novelty (.82), urgency (.77), and power (.77), but considerably lower for the remaining dimensions (.58 – .62).

Acoustic cue	Novelty	Pleasantness	Goal conduciveness	Urgency	Power	Norm compatibility
F0 mean	0.44 ***	0.01 ns	0.01 ns	0.39 ***	-0.18 **	-0.02 ns
F0 sd	0.31 ***	0.16 **	0.17 **	0.16 **	0.21 ***	0.13 *
F0 Q1	0.27 ***	-0.04 ns	-0.06 ns	0.27 ***	-0.28 ***	-0.07 ns
F0 Q5	0.43 ***	0.11 ns	0.09 ns	0.32 ***	-0.05 ns	-0.07 ns
F0 fracrise	0.29 ***	0.11 ns	0.12 *	0.23 ***	-0.04 ns	0.07 ns
Intensity M	0.69 ***	0.11 ns	0.11 ns	0.58 ***	0.26 ***	0.08 ns
H1-A3	-0.26 ***	-0.17 **	-0.17 **	-0.14 *	-0.21 ***	-0.14 *
HF-500	0.34 ***	0.14 *	0.13 *	0.23 ***	0.31 ***	0.15 *
F1 bw	-0.10 ns	-0.18 **	-0.21 ***	-0.03 ns	-0.20 ***	-0.23 ***
% silence	-0.26 ***	-0.37 ***	-0.33 ***	-0.16 **	-0.33 ***	-0.34 ***
Duration	-0.37 ***	-0.10 ns	-0.9 ns	-0.37 ***	-0.14 *	-0.04 ns

Discussion

- The results showed that listeners could reliably infer several aspects of emotion-eliciting situations from vocal expressions
- This suggests that vocal affect expressions may carry cognitive representational information (or *symbolic function*, using Bühler's (1934) terminology)
- The direct assessment of the amount of information about emotion-eliciting situations that can be perceived from emotion expressions presents a novel way of studying social cognition that is not tied to particular emotion categories

Discussion

- More research is needed to determine which appraisal dimensions can be conveyed through the voice, and what combinations of appraisal dimensions that are best suited for annotating emotional speech corpora
- Future research should also examine similarities and differences between the acoustical correlates of emotion appraisal dimensions and other dimensional representations of emotion (e.g., activation, valence, potency)

Future directions

- Within- and cross-cultural annotation of the VENEC corpus using various response formats (e.g., appraisal ratings, forced-choice, free responses)
- Automatic classification experiments in which we investigate the impact of within- and cross-cultural variation in expressive styles on the acoustics characteristics of various emotions

Acknowledgments

- The research was funded by the Swedish Research Council and Sasakawa Young Leaders Fellowship Fund through grants to PL, and the National Science Foundation through grants to HAE.
- We would like to thank all the numerous people all over the world who have helped us with the VENEC studies so far!